

Mining Google+

- Google+ is a social media platform that allows its users to create profile pages and submit posts to interest-based communities.
- It's become a useful platform for data mining.
- Google+ initially serves as our primary source of data for this chapter because it's inherently social, features content that's often expressed as longer-form notes that resemble blog entries, and is now an established staple in the social web.
- It's not hard to make a case that Google+ activities (notes) fill a niche somewhere between Twitter and blogs.
- The concepts from previous chapters could equally be applied to the data behind Google+'s diverse and powerful API, although we'll opt to leave regurgitations of those examples (mostly) as exercises for the motivated reader.
- Wherever possible we won't reinvent the wheel and implement analysis tools from scratch, but we will take a couple of "deep dives" when particularly foundational topics come up that are essential to an understanding of text mining.
- Its rich suites of APIs can be a bit overwhelming at first, but don't worry: while text analytics is an incredibly diverse and complex field of study, there are lots of powerful fundamentals that can take you a long way without too significant of an investment.